Intl. J. River Basin Management Vol. 6, No. 3 (2008), pp. 233–241

A linear geospatial streamflow modeling system for data sparse environments

KWABENA O. ASANTE*, GULEID A. ARTAN†, SHAHRIAR PERVEZ and JAMES ROWLAND, *Science Applications International Corporation (SAIC), contractor to U.S. Geological Survey (USGS) Center for Earth Resources Observation and Science, Sioux Falls, SD 57198, U.S.A. Work performed under USGS contract 03CRCN0001. E-mails: *asante@usgs.gov; †gartan@usgs.gov (†author for correspondence)*

ABSTRACT
In many river basins around the world, inaccessibility of flow data is a major obstacle to water resource studies and operational monitoring. This paper describes a geospatial streamflow modeling system which is parameterized with global terrain, soils and land cover data and run operationally with satellite-derived precipitation and evapotranspiration datasets. Simple linear methods transfer water through the subsurface, overland and river flow phases, and the resulting flows are expressed in terms of standard deviations from mean annual flow. In sample applications, the modeling system was used to simulate flow variations in the Congo, Niger, Nile, Zambezi, Orange and Lake Chad basins between 1998 and 2005, and the resulting flows were compared with mean monthly values from the open-access Global River Discharge Database. While the uncalibrated model cannot predict the absolute magnitude of flow, it can quantify flow anomalies in terms of relative departures from mean flow. Most of the severe flood events identified in the flow anomalies were independently verified by the Dartmouth Flood Observatory (DFO) and the Emergency Disaster Database (EM-DAT). Despite its limitations, the modeling system is valuable for rapid characterization of the relative magnitude of flood hazards and seasonal flow changes in data sparse settings.

Keywords: Streamflow modelling; remote sensing; flow anomaly; extreme events; TRMM.

1 Introduction

In situ river gauges have served as the primary method for operational monitoring of streamflow conditions and for obtaining data required for managing water resources, quality and hazards posed by extreme hydrologic events. However, the past two decades have seen serious declines in global streamflow monitoring infrastructure, particularly in the poorest countries in the world, which are most vulnerable to changes in water quantity and hazards characteristics (Stokstad, 1999). Increased frequency of extreme precipitation, temperature and drought events associated with global climate change (Intergovernmental Panel on Climate Change, 2007) could further contribute to the decline in availability of hydrologic information in some parts of the world. Extreme floods destroy river gauging infrastructure while low flows associated with droughts tend to concentrate in small rills away from the river gauges. In light of these limitations, hydrologic modeling is becoming an increasingly important complement to in situ gauges for monitoring streamflow.

Numerous rainfall-runoff models of varying complexity have also been developed for estimating design flows. Most of these models are designed for use in small river basins, and they include hydrologic processes that do not scale up well to large river basins. Other water resource models employing statistical approaches such as neural networks to extract implicit relationships between rainfall and runoff require observed streamflow data representing a wide range of hydrologic conditions for training. The climate modeling community has also developed a series of large-area models such as the Variable Infiltration Capacity (VIC) model of Liang et al. (1994) for use in climate change analysis. Flows generated with input from Global Climate Models (GCMs) are not suitable for short-range, operational monitoring. Full supported operational modeling systems such as the U.S. National Weather Service River Forecast System (Monroe and Anderson, 1974) and the European Flood Forecasting System (De Roo *et al.*, 2003) are also in use in data rich settings.

The challenge in data sparse environments is to develop models that can provide useful information on current streamflow conditions without the requirement for observed streamflow data for model initialization, calibration or training. Progress is being made in the development of methods for streamflow prediction in ungauged basins through initiatives such as the Model Parameter Estimation Experiment (MOPEX) (Schaake *et al.*, 2001), the Prediction in Ungaged Basins (PUB) (Sivapalan *et al.*, 2006), and the Distributed Model Intercomparison Project (DMIP) (Reed *et al.*, 2004). However, there is a present need for hydrologic models which can be set up with currently available datasets for streamflow applications in data sparse settings. In this paper, we present the development of a geospatial streamflow modeling system which is capable of producing current streamflow information from globally available initialization and forcing datasets.

Received on September 18, 2007. Accepted on January 23, 2008.

1.1 Study sites

The six largest basins in Africa, namely the Congo (3,731,000 km^2), the Niger (2,262,000 km^2), the Nile (3,255,000 km^2), the Zambezi (1,332,000 km^2), the Orange (941,000 km^2) and Lake Chad (2,498,000 km^2), were selected as study sites. Each of the basins has a completely different set of climatic and geophysical characteristics, providing the variety necessary to test the performance of a large-area hydrologic model. The Congo basin, which has the second largest discharges of any river in the world, is almost entirely covered by forests. The Nile, one of the longest rivers in the world, traverses a range of climatic conditions ranging from tropical and upland forests to Mediterranean, semiarid and desert conditions. The Niger originates in the Guinean highlands, makes it way through the Inner Delta in Mali and flows through the arid regions of the Sahel before joining the Benue River on its way to the tropical forests and finally the delta outlet near Port Harcourt in Nigeria. Unlike the other river systems, the Chad basin is an inland catchment which terminates in a series of inland lakes. Many of the lakes are currently not linked due to long-term drought and over utilization of water resources. Each lake could consequently be considered as an independent terminal sink for delineation purposes. However, a single terminal sink, located in Lake Chad, is adopted in this study to ensure consistency of basin boundaries with prior delineations, such as those of Revenga *et al.* (1998) and Olivera *et al.* (2000).

2 The Geospatial streamflow modeling system

The modeling system consists of the stand-alone Geospatial Streamflow Model (GeoSFM) software (Asante *et al.*, 2007a), a baseline parameterization service and an operational data processing service which support end user applications as shown in Figure 1. The GeoSFM software is a semi-distributed hydrologic model developed as an extension of the ArcView Geographic Information System (GIS) software. It contains GIS-based pre-processing and post-processing modules and a hydrologic routing module that uses dynamically linked libraries (DLLs), created in a mixed-programming environment, to perform hydrologic computations and time series manipulation. The U.S. Geological Survey (USGS) Center for Earth Resources Observation and Science (EROS), where the modeling system was developed, maintains the operational data services.

The first pre-processing step is a terrain analysis undertaken to subdivide the study area into catchments and river modeling units and to extract terrain-dependent parameters from a Digital Elevation Model (DEM). The HYDRO1K global elevation dataset (Verdin and Greenlee, 1996) is distributed as the standard dataset for baseline parameterization. Other higher resolution elevation datasets such as the Shuttle Radar Topography Mission (SRTM) data, which are also available from the USGS EROS site, can also be used in GeoSFM. For this study, a terrain analysis was undertaken with a minimum contributing area threshold of 5,000 km^2 for stream network initiation. For the six study basins, the analysis generated between 120 and 385 catchments, with an average area of 8,750 km^2 and an associated river reach about 110 km in length. The catchment and river network delineations for the Chad, Congo, Niger, Nile, Orange and Zambezi basins are shown in Figure 2. Other parameters such as flow direction, upstream drainage area, distance to the nearest stream channel, distance to the basin outlet, and the downstream river reach number are also derived using automated GIS routines in GeoSFM.

The second pre-processing task in the GeoSFM model is the parameterization of the hydrologic modeling units delineated in the first pre-processing step. The terrain analysis parameters described in the preceding paragraph are aggregated to obtain a single value of each parameter for each catchment. Additionally,

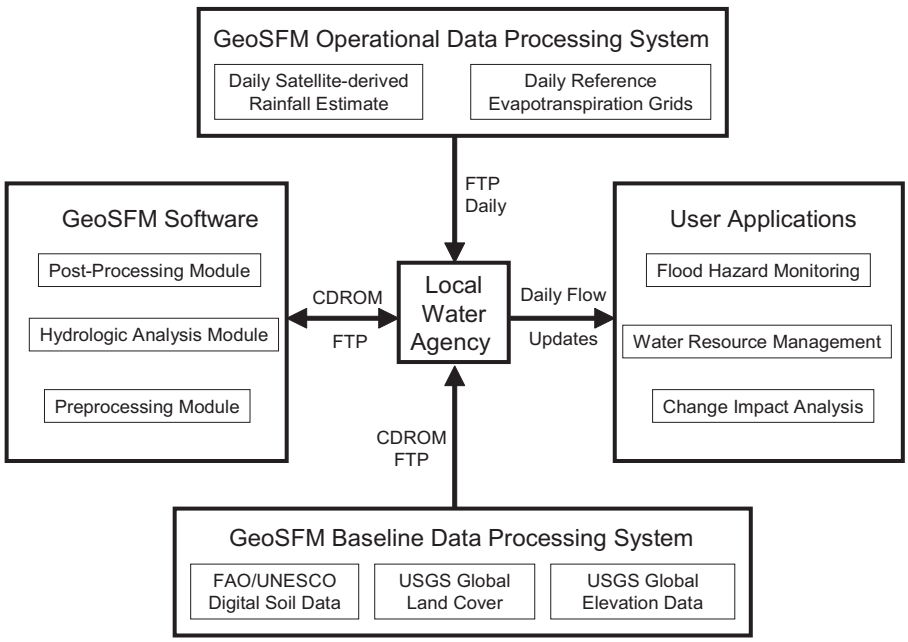

Figure 1 Components of the Geospatial Streamflow Modeling System.

Figure 2 Map of Africa showing the Congo, Niger, Nile, Zambezi, Orange and Lake Chad basins and river networks delineated. The location of stream gauges used in the study are also shown.

remotely sensed vegetation from the Global Land Cover Characteristics (GLCC) database (Loveland *et al.*, 2000) and digital soil data from the Digital Soil Map of the World (Food and Agriculture Organization, 1998) and the World Soil File (Zobler, 1986) are introduced to determine the predominant land cover and soil texture classes in each catchment. Based on these classes, published values of Manning roughness coefficient (McCuen, 1998) and soil parameters such as water holding capacity and hydraulic conductivity (Webb *et al.*, 1993) are assigned to each catchment. Dispersion coefficient for the diffusion analog equation is computed from the Muskingum X assigned to each river reach by equating the second moments (Dooge, 1973) of the response functions for the Muskingum and diffusion analog equations. Since values of Muskingum X can be estimated using rules of thumb (McCuen, 1998), corresponding values of the dispersion coefficient can be obtained for each river reach of known length and celerity.

2.1 Satellite-derived fluxes

The primary input fluxes for GeoSFM are satellite-derived precipitation and evapotranspiration. The operational data processing system supports the daily processing and distribution of these datasets. The Tropical Rainfall Measuring Mission (TRMM) of the U.S. National Aeronautics and Space Administration (NASA) produces merged three-hourly rainfall rates incorporating spaceborne radar, microwave data and infrared imagery. These data are processed at USGS EROS to convert them to daily accumulations and reformatted to GIS-ready images. The NASA TRMM product (version 3B42) covers the tropics between 50° north and 50° south with grid cells of spatial resolution 0.25° by 0.25°. The NASA TRMM products have a daily climatology beginning in 1998 to the present. Operationally, the most current products are made available about 12 hours after the end of the collection period. While other satellite-derived rainfall products are available, the NASA TRMM products are used in this application because of their superior performance in regions with limited situ gauges (Dinku *et al.*, 2007).

The operational data processing system also produces and distributes a daily reference evapotranspiration (ET_o) dataset with global coverage as described in Verdin and Klaver (2002). The dataset is produced by ingesting output fields from NOAA's Global Data Assimilation System (GDAS) (Kanamitsu, 1989) into the Penman-Monteith equation (1). It assumes a hypothetical reference crop with an assumed crop height of 0.12 m, a fixed surface resistance of 70 s m^{-1} and an albedo of 0.23 (Allen *et al.*, 1998). The computation is performed at an hourly interval and aggregated to obtain a daily value of ET_o. However, the input fields remain constant for at least 6 hours during the computation since GDAS data are updated every 6 hours.

$$ET_o = \sum_{t=1}^{24}\left(\frac{1}{(\Delta+\gamma(1+0.34u_2))}\right) * \left[(0.408\Delta\,(R_n-G)) + \left(\frac{37\gamma u_2\,(e_s-e_a)}{(T+273)}\right)\right] \quad (1)$$

where ET_o is the reference evapotranspiration [mm day^{-1}], is the slope of the saturation vapour pressure [kPa °C^{-1}], is the psychrometric constant [kPa °C^{-1}], is the wind speed at 2 m above the ground surface [m s^{-1}], is the net radiation [MJ m^{-2} day^{-1}], is the soil heat flux [MJ m^{-2} day^{-1}], T is the air temperature [°C], is the saturation vapour pressure [kPa °C^{-1}] and is the actual vapour pressure [kPa °C^{-1}].

The radiation and heat fluxes are generated by the GDAS model while the vapour pressure terms (Δ, γ, e_s and e_a) are computed from GDAS temperature and humidity fields. Wind fields computed by GDAS at 10 m heights are downscaled to obtain 2 m wind fields for use in the evapotranspiration computations. The resulting product has a spatial resolution of 1° by 1°.

2.2 Flow routing

In simulating flow, mean areal precipitation and evapotranspiration values for each catchment are determined by spatial averaging of daily rainfall and ET_o grids, and the resulting time series are stored in ASCII files. GeoSFM contains a number of linear and non-linear routines for soil moisture accounting and in-stream flow routing. The simpler linear soil accounting routine is used in this application. It employs variable contributing areas (based on degree of saturation) for surface runoff generation and a linear reservoir for subsurface runoff generation. The equations described in this section are applied to each of the small catchments which have a typical area of approximately 8,750 km^2. For each daily time step, the change in soil moisture storage in each catchment is computed from the continuity equation (2).

$$\frac{\Delta S_j^i}{\Delta t} = P_j^i - E_j^i - R_j^i - G_j^i \quad (2)$$

where P_j^i is the precipitation in mm at time i, E_j^i is the actual evapotranspiration in mm computed as the lower of ET_o and available soil moisture, R_j^i is the total runoff in mm, G_j^i is the deep percolation to ground water in mm and S_j^i is the available soil moisture in mm at time i. The actual evapotranspiration is limited by the lower of the ET_o and the moisture available in soil storage.

The rate of percolation to deep ground water is governed by a linear reservoir with residence time computed as the total soil depth divided by the saturated hydraulic conductivity as shown in equation (3).

$$G_j^i = S_j^i \cdot \frac{Ks_j}{SDEPTH_j} \cdot \exp\left(-\frac{Ks_j}{SDEPTH_j}\right) \tag{3}$$

where G_j^i is the percolation for each day in mm at time i, $SDEPTH_j$ is the soil depth in mm and Ks_j is the saturated hydraulic conductivity in mm/day.

The total runoff is computed as a summation of the excess precipitation and the baseflow generated from soil storage using linear reservoir functions with quick and slow components. If the soil column is saturated, all excess precipitation is converted to runoff. For unsaturated soil conditions, runoff is generated by using partial contributing areas as given by equation (4).

$$R_j^i = P_j^i \cdot f\left(\frac{S_j^i}{SMAX_j}\right) \tag{4}$$

where R_j^i is the runoff for each day in mm at time i, S_j^i is the available soil moisture in mm at time i, $SMAX_j$ is the soil water holding capacity in mm and f is a function defining the relationship between percent soil saturation and percent impervious cover.

A dimensionless unit hydrograph is generated for each catchment by discretizing the flow times for grid cells within the catchment. The unit hydrograph is used to route runoff from the water balance to the catchment outlet where it enters the next downstream river reach as shown in equation (5).

$$q_j^i = 0.001 \cdot A_j R_j^i * \sum_{i=1}^{m} U_j \tag{5}$$

where q_j^i is the overland flow in m^3/s arriving at the catchment outlet j at time i, A_j is the surface area of the runoff generating unit in m^2, U_j is the dimensionless unit hydrograph for catchment j, and $*$ is the convolution integral.

At the catchment outlet, overland flow is added to flow already in the river reach from upstream catchments and routed to the next downstream end of the river reach. The distribution of inflow at the upstream end of each river reach which arrives at the downstream end of the same reach is computed using the diffusion analog equation (6).

$$h_j(t) = \frac{x_j}{\sqrt{4\pi D_j t^3}} \cdot \exp\left[-\frac{(x_j - v_j t)^2}{4 D_j t}\right] \tag{6}$$

where $h_j(t)$ is the fraction of inflow arriving at the downstream end of each river reach, x_j is the length of each river reach in m, v_j is the kinematic wave velocity in m/s, D_j is the dispersion coefficient in m^2/s, and t is the time after input in seconds.

The discharge at the downstream end of each river reach is computed by convolving the diffusion analog response function with flow entering the upstream end (equation 7), and the discharge is passed on as inflow to the next downstream river reach.

$$Q_j^i = h_j(t) * \sum_{k=1}^{n} q_k^i \tag{7}$$

where Q_j^i is the discharge arriving at the outlet of river reach j in m^3/s, q_k^i is the inflow from each river entering the upstream end of reach j in m^3/s, n is the number of river reaches immediately upstream of reach j, and $*$ is the convolution integral. For each simulation time step, daily flow values, in m^3/s, for the downstream end of each of the 100 or more river reaches in the study basin are estimated and written to an output file. However, these flow values contain significant biases which must be taken into consideration when interpreting model results.

2.3 Interpreting simulation results

A number of validation studies have been conducted with TRMM and other satellite-derived rainfall estimates in various part of Africa including the Sahel (Nicholson *et al.*, 2003), the East Africa highlands (Dinku *et al.*, 2007) and zones representing arid, semiarid, savanna and tropical wet climates (Adeyewa and Nakaruma, 2003). These studies indicate the existence of strong regional and seasonal biases attributable to variations in rainfall generation-mechanisms, intensity and frequency. It is therefore necessary to apply bias corrections based on in-situ gauge data before the satellite-derived rainfall estimates can be used to model streamflow quantities (Hughes, 2006; Artan *et al.*, 2007). However, bias correction is not always possible since in-situ gauge data are often not readily available or easily accessible.

To facilitate streamflow monitoring with satellite-derived rainfall in the absence of bias correction, daily streamflow values from GeoSFM simulations are aggregated to monthly intervals and presented as relative anomalies rather than absolute flows. The aggregation reduces the impact of random errors while the use of relative anomalies reduces the impact of systematic biases. Conversion from absolute to relative flows is achieved by comparing modeled flows to similarly modeled flows from preceding time periods. The monthly flows are first converted to anomalies by subtracting the long-term mean flows. The anomalies are then divided by the standard deviation computed from the 12 monthly values in an average year. Computing anomalies this way limits the range of flow anomalies to a few standard deviations, allowing hydrographs from different basins to be compared alongside each other.

3 Sample simulation results

3.1 Flow regime characterization

Characterizing the normal patterns of seasonal flow variation and the predictability of these patterns provides useful information for

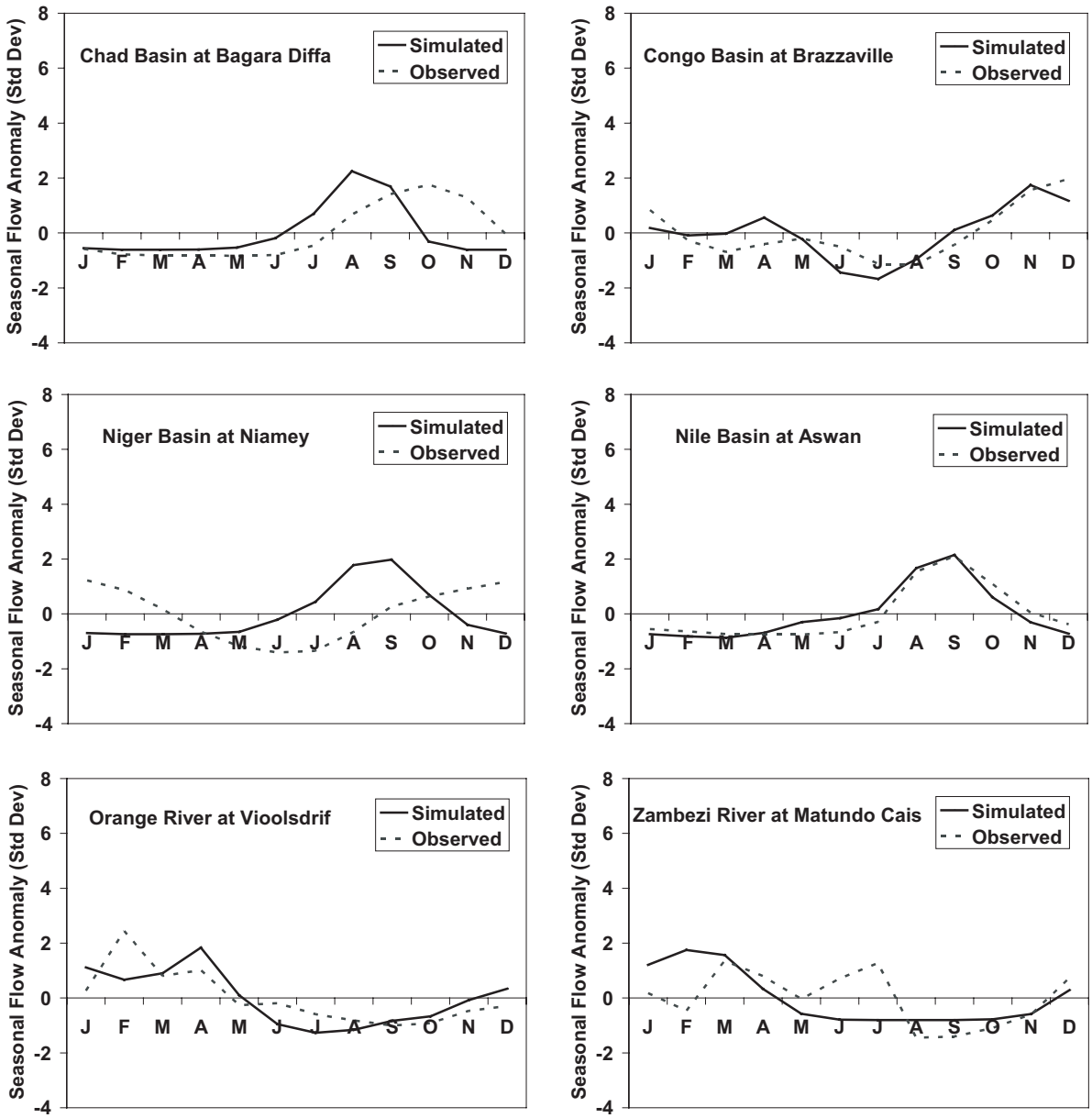

Figure 3 A comparison of characteristic monthly flow hydrographs from GeoSFM simulations and stream gauge data from the Global River Discharge Database. The monthly flows are expressed in number of standard deviations from mean flow.

understanding and managing a river system. The normal patterns also provide a context for comparison of future events or climate-related changes. Figure 3 shows mean monthly flows in the six test basins during the simulation period from January 1998 to December 2005. The dark continuous lines on the graph show flows presented in terms of the standard deviation among monthly values during the mean year. The dark dotted lines show the mean monthly flows from the Global River Discharge Database (RivDIS) of Vörösmarty *et al.* (1996) which contains flows for hydrometric stations around the world. RivDIS includes records spanning the period 1813 to 1991. However, the actual period of record for most stations is significantly shorter with a mean of about 19 years of data per station.

The spatiotemporal distribution of gauged streamflow data is also very uneven. For example, Matundo Cais, the most downstream gauging location on the main stem of the Zambezi, only contains 4 full years of data, while Aswan on the Nile has over 114 years of data available. Other stations used in this comparison include Vioosdrif on the Orange (22 years), Bagara Diffa in the Chad basin (23 years), Brazzaville on the Congo (18 years), and Niamey on the Niger (50 years). The observed flow data predate the satellite-derived precipitation data from TRMM. However, they are used in this application because they are currently the only open-access database of global streamflow available.

The comparisons reveal that the modeling approach captures the relative magnitude of characteristic flows quite accurately in the Nile, Congo, and to a lesser extent the Orange basin. The hydrographs in the Zambezi are generally similar, but the observed flow record is too short to allow for detailed analysis off the differences. In the Niger and Chad basins, the simulated hydrographs peak too early compared to the historical observations. Similar early peaks in simulated hydrographs have been

noted in the Inner Delta in the Niger basin (Ducharne *et al.*, 2003) and in the western marshes in the Chad basin (Coe and Birkett, 2004). These studies have noted a reduction in travel time (as high as 30-days) beginning in 1997. The changes may be attributable to changes in marsh vegetation distribution, channel connectivity and surface-subsurface flow dynamics occurring in inland marshes under the influence of climate variability and change. Nonetheless, these problems highlight the need for improved algorithms for accounting for spatial and temporal variations in flow velocities in different terrain and land cover types in GeoSFM and other large scale hydrologic models.

3.2 Flow anomalies characterization

Identification of severe flood or drought events from GeoSFM simulation results allows for the development of hazard profiles, particularly in areas where data availability, accessibility, or communication problems limit access to hazard warning information (Asante *et al.*, 2007b). These hazard profiles can serve as a basis for designing measures to minimize loss of life and damage to property. Figure 4 shows monthly flows in the six test basins between 1998 and 2005. The flows are presented in terms of standard deviations from the mean flow, computed from all 96 monthly flow values for each basin.

Though their magnitudes vary significantly from year to year, the hydrographs for the two Sahelian basins, the Niger and the Chad, repeat annually with little changes in shape and timing of peak flows. The Zambezi and the Nile also experience minor changes in shape because of interannual differences in the relative magnitudes of runoff from two different sources within each basin. In the case of the Zambezi, the two runoff sources are associated with the passage of the Inter-Tropical Convergence Zone (ITCZ) over the basin during its southward and northward migrations. In the Nile, the two sources of runoff are the non-contiguous Great Lakes region for the White Nile and the Ethiopian Highlands for the Blue Nile. Each year, there are

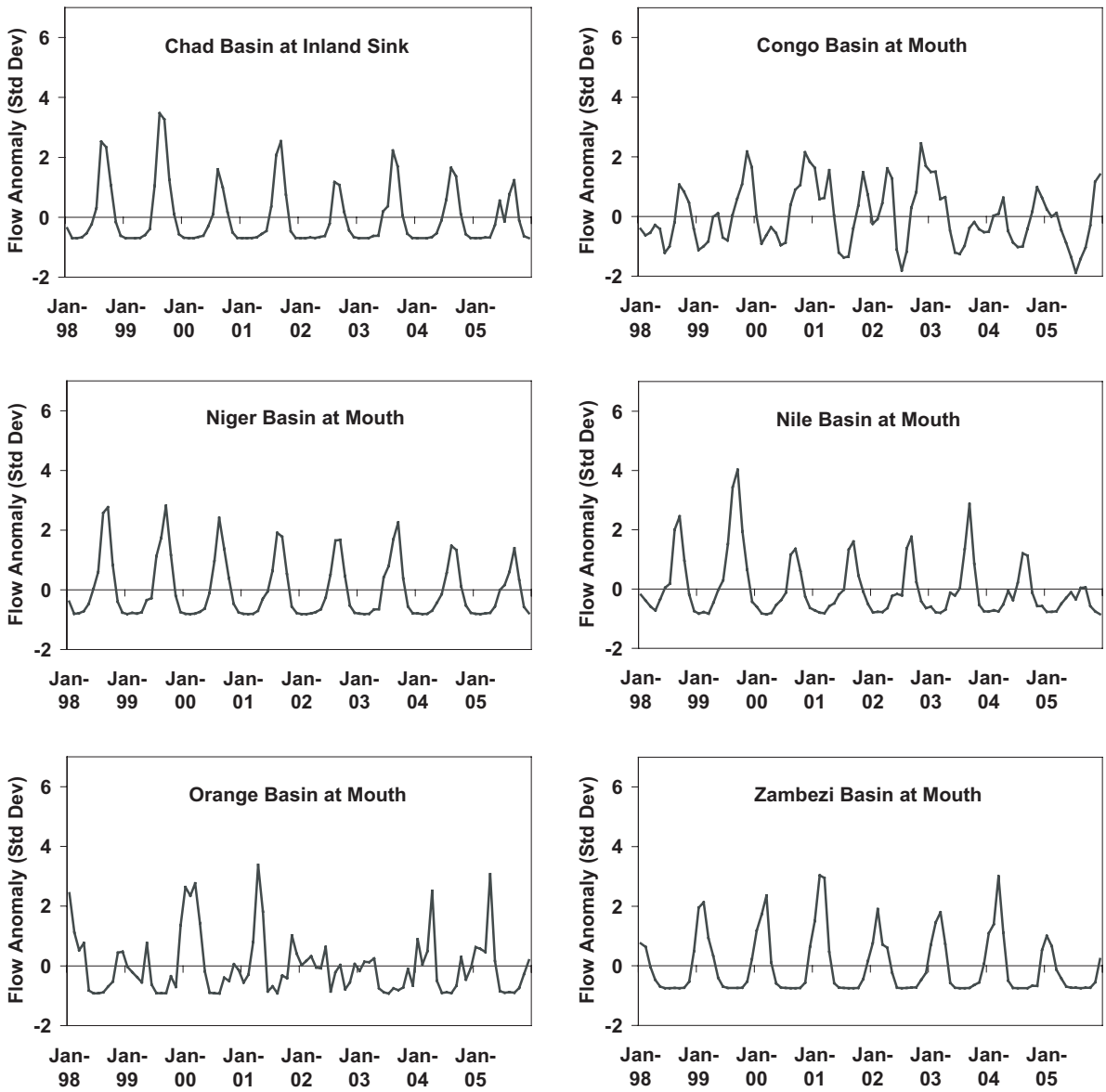

Figure 4 Simulations of monthly flow anomalies discharged at the mouth or inland sink of the Lake Chad, Congo, Niger, Nile, Orange and Zambezi basins between January 1998 and December 2005.

Table 1 Listing of the four wettest periods in each test basin and the independent confirmation of humanitarian impacts in the Emergency Disasters Database and Dartmouth Flood Observatory (DFO) archive of large flood events. The magnitudes of the flood events are presented in terms of standard deviations from mean flow. Both the mean and standard deviation are computed from monthly flows.

Basin	Source	Event 1	Event 2	Event 3	Event 4
Chad	Simulation	3.5 (Aug 1999)	2.5 (Aug 1998)	2.5 (Sep 2001)	2.2 (Aug 2003)
	Verification	EMDAT, DFO	–	EMDAT[1], DFO[2]	DFO[2]
Congo	Simulation	2.5 (Nov 2002)	2.2 (Nov 2000)	2.2 (Nov 1999)	1.6 (Apr 2002)
	Verification	EMDAT[1], DFO[2]	–	EMDAT[1], DFO[2]	–
Niger	Simulation	2.8 (Sep 1999)	2.8 (Sep 1998)	2.4 (Aug 2000)	2.3 (Aug 2003)
	Verification	EMDAT[1], DFO[2]	EMDAT[1], DFO[2]	EMDAT[1]	EMDAT[1], DFO[2]
Nile	Simulation	4.0 (Sep 1999)	2.9 (Sep 2003)	2.5 (Sep 1998)	1.8 (Sep 2002)
	Verification	EMDAT[1], DFO[2]	EMDAT[1]	EMDAT[1], DFO[2]	EMDAT[1]
Orange	Simulation	3.4 (Apr 2001)	3.1 (Apr 2005)	2.8 (Mar 2000)	2.5 (Jan 1998)
	Verification	–	–	–	–
Zambezi	Simulation	3.0 (Feb 2001)	3.0 (Mar 2004)	2.4 (Mar 2000)	2.1 (Feb 1999)
	Verification	EMDAT[1], DFO[2]	DFO[2]	EMDAT[1], DFO[2]	EMDAT[1], DFO[2]

[1]Emergency Disasters Database (EMDAT)
[2]Global Active Archive of Large Flood Events from Dartmouth Flood Observatory (DFO)

differences in the relative magnitude and timing of flows from these sources resulting in slight changes in hydrograph shape. The annual hydrographs in the Congo and the Orange Rivers exhibit much more interannual variability in distribution and timing of flows. Improved monitoring of this variability is extremely important, particularly in the Congo basin which has the potential to influence regional and even global climate through its significant freshwater contribution to the Atlantic Ocean.

3.3 Extreme event verification

The flow anomalies presented in Figure 4 can also be used for identifying extremely wet periods when flooding is likely to occur. Table 1 presents the magnitude (in standard deviations from the mean) of the four wettest months in each basin, with a single peak wetness month selected to represent extended periods of successively wet months. The peaks from the GeoSFM simulation were compared with flood events reported in the global flood archive maintained by the Dartmouth Flood Observatory (DFO) (Brakenridge, 1999) and the Emergency Disaster Database (EM-DAT) (Sapir and Misson, 1992). The DFO archive identifies flood events through media reports with independent verification from satellite imagery. EM-DAT relies on disaster declarations and humanitarian assistance calls as well as media reports of fatalities or impacted populations to initiate a record of an event.

Of the 24 wet months identified in the simulations, 17 were reported in one of the two archives as having resulted in flooding with significant humanitarian impacts. Twelve of the 17 events were reported in both archives. All four events in each of the Niger, Chad, Nile and Zambezi basins were verified as having resulted in humanitarian disasters. This result emphasizes the fact that simulations are useful for identifying flood events, even when aggregated to monthly scales. It also highlights the potential complementary role for operational simulation models in identifying humanitarian events which one of the current global archives would have missed on their own. The remaining six unverified events occurred in the Congo and Orange basins which also exhibited the most variability in shape and timing of annual flows. Several large dams in the Orange basin including Gariep (5341 million m^3), Van der Kloof (3171 million m^3), Vaal (2603 million m^3) and Bloemhof (1264 million m^3) provide adequate storage to absorb most flood events, thus diminishing the likelihood of a humanitarian disaster. This result highlights a potential problem posed by not incorporating reservoir locations and operations in the GeoSFM simulations. A development of a global database of reservoir locations, capacities and operating rules would facilitate such modeling efforts. However, the absence of reservoir data is unlikely to explain the two unverified events in the Congo because of the absolute magnitude of typical flows in the basin. A global database specifying the location and timing of flood hazards from simulated flows would be a useful complement to existing disaster databases.

4 Conclusions

This paper has presented a geospatial streamflow modeling system which is initialized and operated with geospatial datasets of topography, land cover, soils, daily precipitation and evapotranspiration which are freely available on the Internet. Descriptions of soil moisture accounting and flow routing algorithms included in the modeling system have also been provided. Model results have been presented in relative terms to ensure appropriate use of the information. In applications in several large basins in Africa, the system is used to characterize the relative magnitude of mean monthly flows and to identify extreme flood events. The main limitations of GeoSFM including its inability to predict absolute flow magnitude and difficulties in characterizing flow travel time in basins with significant wetlands or reservoir systems.

In spite of these limitations, the information value of GeoSFM's relative flow anomalies is demonstrated by independent verification of most major simulated flood events using records from two open-access flood disaster databases. The model can provide independent monitoring information to water managers working in river systems with limited in-situ data. A common example is in international river basins where administrative problems prevent downstream countries from accessing in-situ information on extreme events originating upstream of the international boundary. Reservoir manager can also use GeoSFM to generate seasonal reservoir inflow anomalies from seasonal rainfall forecasts. Flow anomalies can inform water manager on how expected inflows for the forecast period will compare with prior years. In such applications, GeoSFM-generated seasonal inflow anomalies can provide a basis for making medium range water management decisions. In summary, GeoSFM can fill important information gaps in extreme event monitoring and medium range water resource management in data sparse settings.

Acknowledgements

The authors would like to acknowledge two anonymous reviewers, Gabriel Senay, Bruce Worstell and Thomas Adamson for helpful comments on the manuscript. The authors also acknowledge the Office of Foreign Disaster Assistance of the U.S. Agency for International Development (USAID/OFDA) and the U.S. Geological Survey (USGS) for funding this work.

References

1. Adeyewa, Z.D. and Nakamura, K. (2003). "Validation of TRMM Radar Rainfall Data Over Major Climatic Regions in Africa," *Journal of Applied Meteorology*, 42, 331–347.
2. Artan, G., Gadain, H., Smith, J.L., Asante, K.O., Bandaragoda, C.J. and Verdin J.P. (2007). "Adequacy of Satellite Derived Rainfall Data for Stream Flow Modeling," *Natural Hazards*, 42(3), 167–185.
3. Asante, K.O., Artan, G.A., Pervez, S., Bandaragoda, C. and Verdin, J.P. (2007a). "Technical Manual for the Geospatial Streamflow Model (GeoSFM)," *U.S. Geological Survey Open-File Report 2007–1441*, pp. 1–99.
4. Asante, K.O., Dezanove, R.M., Artan, G.A., Lietzow , R. and Verdin, J.P. (2007b). "Developing a Flood Monitoring System From Remotely Sensed Data for the Limpopo Basin," *Transactions on Geoscience and Remote Sensing*, 45(6), 1709–1714.
5. Allen, R.G., Pereira, L.S., Raes, D. and Smith, M. (1998). Crop Evapotranspiration – Guidelines for Computing Crop Water Requirements – *FAO Irrigation and Drainage Paper 56*, Food and Agriculture Organization of the United Nations, Rome, Italy.
6. Brakenridge, G.R. (1999). "River flooding and Global Climate Change: A Multisensor Approach," in *Earth Science Enterprise Reference Handbook*, R. Greenstone (ed.), EOS Project Science Office, Code 900, NASA/Goddard Space Flight Center, Greenbelt, Maryland, pp. 163–165.
7. De Roo, A.P.J., Gouweleeuw, B., Thielen, J., Bartholmes, J., Bongioannini-Cerlini, P., Todini, E., Bates, P.D., Horritt, M., Hunter, N., Beven, K., Heise, E., Rivin, G., Hils, M., Hollingsworth, A., Holst, B., Kwadijk, J., Reggiani, P., van Dijk, M., Sattler, K. and Sprokkereef, E. (2003). "Development of a European Flood Forecasting System," *International Journal of River Basin Management*, 1(1), 49–59.
8. Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J. and Ropelewski, C. F. (2007). "Validation of Satellite Rainfall Products Over East Africa's Complex Topography," *International Journal of Remote Sensing*, 28(7), 1503–1526.
9. Dooge, J.C. (1973). "Linear Theory of Hydrologic Systems," *USDA Technical Bulletin no. 1468*, USDA, Washington, DC.
10. Ducharne, A., Golaz, C., Leblois, E., Laval, K., Polcher, J., Ledoux, E. and De Marsily, G. (2003). "Development of a High Resolution Runoff Routing Model, Calibration and Application to Assess Runoff from the LMD GCM," *Journal of Hydrology*, 280(1–4), 207–228.
11. Food and Agriculture Organization (FAO), (1998). "Digital Soil Map of the World and Derived Soil Properties," *FAO Land and Water Media Series*, Rome, Italy.
12. Hughes, D.A. (2006). "Comparison of Satellite Rainfall Data with Observations from Gauging Stations Networks," *Journal of Hydrology*, 327, 399–410.
13. Intergovernmental Panel on Climate Change (IPCC), (2007). Climate Change 2007 – The Physical Science Basis, *Contribution of Working Group I to the Fourth Assessment Report of the IPCC*, Cambridge University Press, New York, NY.
14. Kanamitsu, M. (1989). Description of the NMC Global Data Assimilation and Forecast System, *Weather Forecasting*, 4, 335–342.
15. Liang, X., Lettenmaier, D.P., Wood, E.F. and Burges, S.J. (1994). "A Simple Hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs," *Journal of Geophysical Research*, 99(D7), 14415–14428.
16. Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, J., Yang, L. and Merchant, J.W. (2000). "Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data," *International Journal of Remote Sensing*, 21(6), 1303–1330.
17. McCuen, R.H. (1998). *Hydrologic Analysis and Design*, Prentice Hall, Upper Saddle River, NJ.
18. Monroe, J.C. and Anderson, E.A. (1974). "National Weather Service River Forecasting System," *ASCE Journal of Hydraulics Division*, 100(HY5), 621–630.
19. Nicholson, S.E., Some, B., McCollum, J., Nelkin, E., Klotter, D., Berte, Y., Diallo, B.M., Gaye, I., Kpabeba, G., Ndiaye, O., Noukpozounkou, J.N., Tanu, M.M., Thiam, A., Toure, A.A. and Traore, A.K. (2003). "Validation of TRMM and Other Rainfall Estimates with a High-Density Gauge Dataset for West Africa.

Part II: Validation of TRMM Rainfall Products," *Journal of Applied Meteorology*, 42, 1355–1368.

20. OLIVERA, F., FAMIGLIETTI, J.S. and ASANTE, K.O. (2000). "Global-Scale Flow Routing Using a Source to Sink Algorithm," *Water Resources Research*, 36(8).
21. REED, S., KOREN, V., SMITH, M., ZHANG, Z., MOREDA, F., SEO, D.-J. and DMIP Participants (2004). "Overall Distributed Model Intercomparison Project Results," *Journal of Hydrology*, 298(1–4), 27–60.
22. REVENGA, C., MURRAY, S., ABRAMOVITZ, J. and HAMMOND, A. (1998). *Watersheds of the world: Ecological Value and Vulnerability*, World Resources Institute, Washington, DC.
23. SAPIR, D.G. and MISSON, C. (1992). "The Development of a Database on Disasters," *Disasters*, 16(1), 80–86.
24. SCHAAKE, J.C., DUAN, Q. and HALL A. (2001). "Toward Improved Parameter Estimation of Land Surface Hydrology Models through Model Parameter Estimation Experiment (MOPEX)," *IAHS Publ. No. 270*, pp. 91–97.
25. SIVAPALAN, M., WAGENER, T., UHLENBROOK, S., ZEHE, E., LAKSHMI, V., LIANG, X., TACHIKAWA, Y. and KUMAR, P. (eds.) (2006). "Predictions in Ungauged Basins: Promises and Progress," *IAHS Publication 303*.
26. STOKSTAD, E. (1999). "Scarcity of Rain, Stream Gages Threatens Forecasts," *Science*, 285, 1199.
27. VERDIN, J.P. and KLAVER, R.W. (2002). "Grid Cell Based Crop Water Accounting for the Famine Early Warning System," *Hydrologic Processes*, 16, 1617–1630.
28. VERDIN, K.L. and GREENLEE, S.K. (1996). "Development of Continental Scale Digital Elevation Models and Extraction of Hydrographic Features," *Proceedings: 3rd International Conference on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico*, National Center for Geographic Information and Analysis, Santa Barbara, CA.
29. VÖRÖSMARTY, C.J., FEKETE, B. and TUCKER, B.A. (1996). "River Discharge Database (RivDIS), Version 1.0, A Contribution to IHP-V Theme 1," *Technical Documents in Hydrology Series*, UNESCO, Paris.
30. WEBB, R.S., ROSENZWEIG, C.E. and LEVINE, E.R. (1993). "Specifying Land Surface Characteristics in General Circulation Models: Soil Profile Data Set and Derived Water-Holding Capacities," *Global Biogeochemical Cycles*, 7, 97–108.
31. ZOBLER, L. (1986). "A World Soil File for Global Climate Modeling," *NASA Tech. Memo. 87802*, NASA Goddard Institute for Space Studies, New York, NY.